DOCUMENT RESUME

ED 225 473 HE 015 833

AUTHOR	Overall, Jesse U., IV; Marsh, Herbert W.
TITLE	Students' Evaluations of Teaching: An Update. AAHE-ERIC Higher Education Research Currents.
INSTITUTION	American Association for Higher Education, Washington, D.C.; ERIC Clearinghouse on Higher Education, Washington, D.C.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	Dec 82
NOTE	5p.
AVAILABLE FROM	Publications Department, American Association for Higher Education, One Dupont Circle, Suite 780, Washington, DC 20036
PUB TYPE	Information Analyses - ERIC Information Analysis Products (071) -- Collected Works - Serials (022)
JOURNAL CIT	AAHE Bulletin; Dec 1982 p9-12
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	College Faculty; Educational Research; Evaluation Criteria; *Evaluation Methods; *Faculty Evaluation; Higher Education; Institutional Research; Questionnaires; Reliability; *Student Evaluation of Teacher Performance; Teacher Characteristics; *Teacher Effectiveness; *Teaching Skills; Validity

ABSTRACT

Recent research (1978-1982) on student evaluations of teaching is reviewed, including: influence of background variables pertaining to the student, the teacher, and the learning environment; the dimensions of the teaching being evaluated; the validity of students' evaluations; the "Doctor Fox" effect and its implications for validity; the reliability, stability, generalizability, and usefulness of students' evaluations; and the construction and selection of evaluation questionnaires. Dimensions of teaching that students evaluate may include: skill, rapport, assignments, breadth of coverage, tests and grading, group interaction, enthusiasm, and organization. The extent to which students' evaluations of faculty correlate with variables thought to reflect effective teaching may be addressed by considering the following criteria: student achievement, instructor self-evaluations, and improved student attitudes toward the subject. A number of studies have examined the "Doctor Fox" effect: the possibility that student assessments of teacher effectiveness are more a function of an instructor's wit and personality than of the educational content of the lecture. It is concluded that the research indicates that (1) evaluations are not significantly influenced by background variables, and are valid, reliable, stable, generalizable, and useful, and (2) properly constructed evaluation questionnaires assess multiple dimensions of the instructional process. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

AAHE-ERIC HIGHER EDUCATION

Research Currents

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED225473

Students' Evaluations of Teaching: An Update

by Jesse U. Overall, IV and Herbert W. Marsh

Research Currents is prepared by the ERIC Clearinghouse on Higher Education, The George Washington University, Washington, D.C. The material in this publication was prepared pursuant to a contract with the National Institute of Education, U.S. Department of Education. Contractors undertaking such projects under government sponsorship are encouraged to express freely their judgment in professional and technical matters. Prior to publication, the manuscript was submitted to the American Association for Higher Education for critical review and determination of professional competence. This publication has met such standards. Points of view or opinion, however, do not necessarily represent the official view or opinions of either AAHE or the National Institute of Education.

As teachers, administrators, and researchers in higher education with a special interest in student evaluations of effective teaching, we are often asked how one can identify studies relevant to a particular topic in this area from the large number of published student rating studies, and what the most recent research indicates. These are questions that involve all of us in higher education and hence merit an updated analysis.

There is an extensive and still expanding literature base that examines students' evaluations.[1] With few exceptions, current research is defined as that reported in the five-year period of 1978 to the present. While some excellent research is thus not specifically cited, it is referred to in the studies discussed below. Serious researchers are urged to pursue those primary sources on their own.

There are several major areas of consideration in any discussion of student ratings of courses and instructors: influence of background variables; the dimensions of the teaching being evaluated; the validity of students' evaluations; the "Doctor Fox" effect and its implications for validity; and the reliability, stability, generalizability, and usefulness of students' evaluations and the construction and selection of evaluation questionnaires.

[1]We recommend consulting the excellent general review articles (Aubrecht, 1981; Marsh, in press; and, McKeachie, 1979) or books (Centra, 1981; and Millman, 1981) recently published on the subject. These sources have updated and expanded on the much cited and influential research summaries by Costin, Greenough, and Menges (1971) and Kulik and McKeachie (1975). They provide a collection of research findings and outline individual studies whose findings contributed to the formation of research thought predominant in that area.

Should an area of interest be so specialized that it is not covered in one of the review sources listed above, or should the researcher wish a more comprehensive list of studies about a particular topic, the Educational Resources Information Center (ERIC) collection can be helpful. A search of the ERIC holdings will identify relevant articles from most journals plus important conference papers and institutional studies not otherwise published.

Jesse U. Overall is manager of evaluation and personnel research, Office of Institutional Studies, at the University of Southern California.
Herbert W. Marsh is senior lecturer, Department of Education, University of Sydney, Sydney, Australia.

Background Variables

To what extent have selected variables in the teaching and learning environment been found to be associated with student ratings? A great deal of research has focused on the extent to which background variables such as class size, expected grade, reasons for taking a course, the instructor's research productivity, and both students' and instructors' personalities are related to students' evaluations of their courses and instructors. Most of this research has dealt with single background variables or combinations of one or two variables. These approaches produce interesting but piecemeal conclusions, some of which are discussed further.

Background variables under consideration include the following: administrative, course, instructor, and student. In addition, three multivariate studies that investigated the relationship between many of these variables and instructional effectiveness in a single setting are included.

Administrative Variables. Evaluations appear to be somewhat higher if the student evaluator is identified or if the instructor is present when the evaluations are completed (Feldman, 1979). Also, if students believe that their evaluations will influence decisions on promotions, they tend to rate their instructors higher than if they believe their ratings will be used solely for feedback or instructional improvement purposes (McKeachie, 1979).

Course Variables. Classes with very small numbers of students (1 to approximately 30) or very large classes (approximately 100 or more) tend to receive higher evaluations than those with enrollments between these fig-

ures (Aleamoni, 1981; Aubrecht, 1979; Centra, 1981; and, Feldman, 1978). Courses in the humanities receive slightly higher ratings of overall effectiveness when compared with those in the social sciences or natural sciences. Other average differences are found by discipline, depending on the particular evaluation dimension under analysis (Centra, 1981; Feldman, 1978).

Instructor Variables. No specific personality characteristics in an instructor are consistently or significantly related to receipt of a high evaluation from students (McKeachie, 1979). Furthermore, no definite relationship has been found between an instructor's rank and his or her evaluations, with the exception of teaching assistants (TAs). Students tend to rate TAs lower than faculty (Centra, 1981; Marsh, 1980).

Students' knowledge of an instructor's reputation and research productivity apparently has some relationship to their final evaluations, but not to their achievement (Perry, Abrami, Leventhal, and Check, 1979). The majority of studies have found small, insignificant, or no relationships between an instructor's research productivity (numbers of books and articles published) and magnitude—a large number on the positive or negative side—of student ratings (Aleamoni, 1981; Centra, in press).

Researchers have also considered an instructor's teaching experience and teaching load. According to Centra's 1981 summary of research, teacher evaluations tend to improve in the first few years but tend to decline after about 12 years. Centra found no evidence to indicate that evaluations were lower for faculty with larger teaching loads; in fact, he found the opposite result in some circumstances.

Student Variables. Age appears to have little relationship to the magnitude of students' evaluations (Centra, 1981; McKeachie, 1979). There is some evidence that when the gender of both student and instructor is the same, higher evaluations may result on some teaching dimensions (Aleamoni, 1981; Centra, 1981; McKeachie, 1979). Also, neither personality (Abrami, Perry and Leventhal, 1982) nor student/instructor attitude similarity (see Mizener and Abrami, 1981) appears to have any systematic relationship with student ratings. Finally, Aleamoni (1981), Centra (1981) and Feldman (1978) reviewed research and found that while students taking elective courses tend to provide higher ratings, there is no statistically significant relationship between a student's major and his or her rating of a course or instructor.

Multivariate Studies. Recently, three studies were reported in which many of these variables were simultaneously related to instructional effectiveness. Stumpf, Freedman and Aguanno (1979) took the average class ratings given to 129 instructors in all courses taught over two semesters and studied their relationships to several important background variables. Their results indicate only a minor relationship between average ratings and background variables, suggesting that these external factors do not unduly influence the ratings.

Marsh (1980) examined the relationship between students' evaluations and a set of 16 background variables. His multivariate analyses indicate that only 14 percent or less of the variance in the ratings on nine individual rating dimensions or two overall summary items can be explained by the entire set of 16 background variables. The most influential variables were Prior Subject Interest, Expected Grade, course Workload/Difficulty, and percentage of students taking the course for General Interest. In a subsequent analysis (Marsh and Cooper, 1981), Prior Subject Interest was found to account for the largest proportion of variance in students' evaluations (5.1%). This variable also accounted for about one-third of the relationship between ratings and expected grades.

Dimensions of Teaching

What aspects of teaching are students actually evaluating? Common sense and a growing body of empirical research indicate that properly constructed evaluation questionnaires can provide data on several dimensions of teaching. However, it is important to establish the existence of these dimensions through factor analysis before assuming *a priori* that a group of rating items necessarily reflects them.[2]

Frey (1978) analyzed student ratings of instruction and presented solid evidence for two factors or dimensions—Skill and Rapport. Marsh (in press), based on his work with the Student Evaluation of Educational Quality (SEEQ) questionnaire both in this country and in Australia, has identified the following dimensions: Assignments/Readings, Breadth of Coverage, Examinations/Grading, Group Interaction, Individual Rapport, Instructor Enthusiasm, Learning/Value, Organization/Clarity, and Workload/Difficulty. Most of the research in this area, as exemplified by the findings of Frey and Marsh, indicate that properly constructed questionnaires will reflect this multi-dimensionality.

Validity

To what extent do students' evaluations of faculty correlate with variables thought to reflect effective teaching? Because there is no single criterion of effective teaching, several criteria have been selected by consensus as being indicative of instructional effectiveness. Using this construct validation approach, the following criteria appear most relevant.

Achievement. Recent research reported by Centra (1977) and Marsh and Overall (1980) has found positive, significant correlations between average class achievement and magnitude of end-of-term evaluations. Cohen's 1981 review of 41 separate multi-section validity studies reported that student achievement correlated 0.43 with the overall instructor rating and 0.47 with the overall course rating. Mean correlations of 0.22 or higher were noted for all typical rating dimensions except Course Difficulty.

Instructor Self-Evaluations. Braskamp, Caulley and Costin (1979), Doyle and Crichton (1978), and, Marsh, Overall and Kesler (1979) studied the relationship between students' evaluations of individual faculty and the self-ratings of these faculty. In every case, significant positive relationships were found, indicating that instructors and students tended to agree on the effectiveness of instruction in a variety of situations.

Improved Attitudes Toward the Subject. Focusing on an area of increasing interest to researchers, Marsh and

[2]Anyone wishing to develop a questionnaire independently should first read Abrami, Leventhan and Dicken's excellent 1981 discussion of the multi-dimensionality issue. Should one wish to review available instruments that have undergone extensive testing, Braskamp's appendix in Centra's 1981 book provides summary information about the most widely used questionnaires.

Overall (1980) investigated the relationship between students' evaluations and their reported changes in subject matter interest. Positive, statistically significant correlations were found between end-of-term attitudes toward the subject and evaluations received by instructors on most rating dimensions.

The "Doctor Fox" Effect

To what extent can students be lured into providing higher ratings, regardless of lecture content? What are the implications of this research with respect to the validity of students' evaluations? In 1973, Naftulin, Ware and Donnelly reported results from a study suggesting that student assessments of teaching effectiveness were more a function of an instructor's wit and personality than of the educational content of his or her lecture. This investigation and subsequent related research, has been called the "Doctor Fox" studies. The name is based on the pseudonym used by the instructor—a professional actor—who presented the initial lecture.

Further research on this effect by Naftulin, Ware and Donnelly concentrated on the use of video tapes covering six lectures presented by one professional actor. Instructor expressiveness and lecture content were systematically manipulated in each of these lectures in an attempt to replicate the Doctor Fox effect. A review of these studies by Ware and Williams (1979, 1980) led them to conclude that differences in content consistently explains much less variance in students' overall evaluations than do differences in expressiveness.

A 1982 reanalysis of data from the Ware and Williams studies by Marsh and Ware focused on five specific teaching dimensions: Clarity/Organization, Concern, Enthusiasm, Knowledge, and ability to stimulate Learning. Among other findings, they report that in manipulating instructor expressiveness, only ratings on the Enthusiasm dimension were affected; in manipulating the content coverage, only ratings on the Knowledge dimension were affected. This research indicates the importance of using individual evaluation dimensions in addition to overall summary items. It also demonstrates that even if a variable affects instructor ratings, the ratings are not necessarily invalid.

Abrami, Leventhal and Perry (1982) also reviewed the Doctor Fox research. They found that expressiveness has a much larger effect on ratings than on achievement, that content has a much larger effect on achievement than on ratings, and that for either ratings or achievement, the effect of content does not vary to an important extent over levels of expressiveness. They conclude that their findings on the validity of students' evaluations—and similar ones in previous research—can be used as evidence only by first documenting the importance of expressiveness to instruction. They reasoned that because such documentation is absent from earlier research, the Doctor Fox studies—viewed as a measure of instructional processes at work in the field—tell little about the validity or invalidity of students' evaluations.

Reliability/Stability

Do students within the same class agree on the effectiveness of their instructors? Are student ratings stable over time? Research on the reliability of students' evaluations has focused either on the extent to which there is agreement among different students evaluating the same course and instructor (interrater agreement), or agreement among different items purported to measure the same trait (internal consistency). Feldman's 1977 review of this research found that single-rater reliability, when based on a class size of about 20, has reliability coefficients generally greater than 0.80. He thus found the reliability of students' evaluations to compare favorably with the best objective tests if the evaluations are based on a sufficient number of student responses.

In discussions of the stability of students' evaluations it is not uncommon to hear the proposal that retrospective (follow-up) ratings from graduates should be used as the basis for instructor evaluation rather than end-of-term assessments provided by continuing students. The rationale for this view is based on the suggestion that follow-up ratings will in some ways differ significantly from end-of-term ratings because they allow former students to develop additional perspectives about, and to obtain emotional separation from the person and situation being assessed. These follow-up ratings would thus be based on more informed, reflective, and mature judgments.

Overall and Marsh (1980) compare ratings that individual students provided at the end of their courses with subsequent ratings collected from the same students a minimum of one year after graduation. They note insignificant absolute differences between the two sets of ratings, and find a median correlation of 0.83 for all rating dimensions. They conclude that there were no practical differences in the information provided by either set of evaluations, and that these evaluations appear to be quite stable over time.

Generalizability

Are some courses rated less favorably than others on a systematic basis? What is the relative importance of the particular instructor and the specific course in determining student ratings? Research conducted by Gillmore, Kane and Naccarato (1978) led them to conclude that with the instructor as the object of measurement, moderately dependable results can be obtained by generalizing over rating items and students. Furthermore, they conclude that the specific course is not a major factor in determining course evaluations.

Marsh and Overall (1981) studied the relative contribution of the instructor, course level (graduate/undergraduate) and course type (nonquantitative/quantitative) to variance in end-of-term and retrospective student evaluations. They found that the individual instructor performance accounts for 5 to 10 times as much variance in both sets of ratings as did course level or type, suggesting that the particular course subject matter has little effect on student ratings and that the same instructor would probably receive similar ratings in a different course.

In a subsequent study, Marsh (1981) used path analysis to demonstrate that the instructor is the most important determinant of student ratings. Relative to the instructor, the particular course being taught plays a small role.

The results of these studies show consistently that the instructor is more important than the course being evaluated. Thus, ratings obtained by an instructor in one type of course do not necessarily put him or her at a dis-

advantage when one instructor's evaluations are compared with evaluations of instructors teaching other types of courses.

Usefulness

To what extent is feedback from students' evaluations of faculty associated with instructional improvement? Many researchers are concerned with the question of whether instructors who receive summaries of evaluations from their students tend to become more effective teachers and whether students' evaluations are an effective source of feedback in the instructional improvement process. While early research examined by McKeachie (1979) and Rotem and Glasman (1979) found results inconclusive, more recent studies indicate a positive association. Previous research utilized printed feedback provided to each instructor at the end of the term. More recent studies have focused on the addition of individual peer consultation to assist instructors with interpreting and utilizing results from the printed feedback.

Overall and Marsh (1979) found that instructors who receive written feedback from their students at midterm also received more favorable ratings at the end of the term. Their students earned higher final examination scores and reported more favorable affective outcomes. The key ingredient here, they conclude, is the use of an external consultant to interpret the written summaries along the lines suggested by McKeachie and Lin (1975). In a subsequent study, McKeachie et al. (1980) obtained similar findings. They conclude that presentation of encouragement and suggestions to an instructor, in addition to printed feedback, results in a more effective approach to instructional improvement.

Cohen's 1980 meta analysis of feedback studies contains an excellent summary of research in this area. Reviewing 22 college-level studies concerned with this topic, Cohen found a general accentuation of student rating feedback effects when printed summaries were augmented by consultation. These effects were more pronounced for some rating dimensions than for others. Cohen also noted that the positive impact of feedback is not dependent on whether the feedback is used in a midterm/end-of-term or a term-by-term sequence.

Conclusions

A review of recent research concerned with student ratings of teaching indicates that such evaluations are not significantly influenced by background variables, and are valid, reliable, stable, generalizable, and useful. This research further indicates that properly constructed evaluation questionnaires assess multiple dimensions of the instructional process. Thus, when obtained from properly designed questionnaires, data from students' evaluations is particularly appropriate and useful in the instructional evaluation process. ■

Bibliography

Documents available through the ERIC system appear with an ED number. To order, write to ERIC Document Reproduction Service, P.O. Box 190, Arlington, Va. 22210. ED numbers must be specified. All orders must be in writing and include payment. Documents are available as indicated in microfiche (MF) and paper copy (PC).

Abrami, Philip C; Leventhal, Les; and Dickens, Wenda J. 1981. "Multidimensionality of Student Ratings of Instruction." *Instructional Evaluation* 6: 12–17.

Abrami, Philip C.; Leventhal, Les; and Perry, Raymond P. Fall 1982. "Educational Seduction." *Review of Educational Research* 52: 446–464.

Abrami, Philip C.; Perry, Raymond P.; and Leventhal, Les. February 1982. "The Relationship between Student Personality Characteristics, Teacher Ratings, and Student Achievement." *Journal of Educational Psychology* 74: 111–125.

Aleamoni, Lawrence M. 1981. "Student Ratings of Instruction." In *Handbook of Teacher Evaluation*, ed. by Jason Millman. Beverly Hills, California: Sage Publications.

Aubrecht, Judith D. 1981. "Reliability, Validity and Generalizability of Student Ratings of Instruction." *IDEA Paper No. 6*. Kansas State University: Center for Faculty Evaluation and Development. ED 213 296. MF-$1.17; PC-$3.64.

_____. 1979. "Are Student Ratings of Teacher Effectiveness Valid?" *IDEA Paper No. 2*. Kansas State University: Center for Faculty Evaluation and Development. ED 202 410. MC-$1.17; PC not available EDRS.

Branskamp, Larry A.; Caulley, Darrel; and Costin, Frank. Summer 1979. "Student Ratings and Instructor Self-Ratings and Their Relationship to Student Achievement." *American Educational Research Journal* 16: 295–306.

Centra, John A. In press. "Research Productivity and Teaching Effectiveness." *Research in Higher Education*.

_____. 1981. *Determining Faculty Effectiveness*. San Francisco: Jossey-Bass, Inc.

_____. Winter 1977. "Student Ratings of Instruction and Their Relationship to Student Learning." *American Educational Research Journal* 14: 17–24.

Cohen, Peter A. Fall 1981. "Student Ratings of Instruction and Student Achievement: A Meta-Analysis of Multi-Section Validity Studies." *Review of Educational Research* 51: 281–309.

_____. "Effectiveness of Student Rating Feedback for Improving College Instruction: A Meta-Analysis of Findings." *Research in Higher Education* 13: 321–341.

Costin, Frank; Greenough, William T.; and Menges, Robert J. December 1971. "Student Ratings of College Teaching: Reliability, Validity and Usefulness." *Review of Educational Research* 41: 511–536.

Doyle, Kenneth O. and Crichton, Leslie I. December 1978. "Student, Peer, and Self-Evaluations of College Instructors." *Journal of Educational Psychology* 70: 815–826.

Feldman, Kenneth A. April 1979. "The Significance of Circumstances for College Students' Ratings of Their Teachers and Courses." *Research in Higher Education* 10: 149–172.

_____. 1978. "Course Characteristics and College Students' Ratings of Their Teachers: What We Know and What We Don't." *Research in Higher Education* 9: 199–242.

_____. May 1977. "Consistency and Variability among College Students in Rating Their Teachers and Courses." *Research in Higher Education* 6: 223–274.

Frey, Peter W. 1978. "A Two-Dimensional Analysis of Student Ratings of Instruction." *Research in Higher Education* 9: 69–91.

Gilmore, Gerald M.; Kane, Michael T.; and Naccarato, Richard W. Spring 1978. "The Generalizability of Student Ratings of Instruction: Estimates of Teacher and Course Components." *Journal of Educational Measurement* 15: 1–13.

Kulik, James A. and McKeachie, Wilbert J. 1975. "The Evaluation of Teachers in Higher Education." In *Review of Research in Education* (Volume III), ed. Fred N. Kerlinger, Itasca, Illinois: Peacock.

Marsh, Herbert W. In press. "Students as Evaluators of Teaching." *International Encyclopedia of Education: Research and Studies*, ed. by Torsten Husen and T. Neville Postlethwaite. Oxford: Pergamon Press.

_____. February 1982. "SEEQ: A Reliable, Valid, and Useful Instrument for Collecting Students' Evaluations of University Teaching." *British Journal of Educational Psychology* 52: 77–95.

_____. Winter 1981. "The Use of Path Analysis to Estimate Teacher and Course Effects in Student Ratings of Instructional Effectiveness." *Applied Psychological Measurement* 6: 47–60.

_____. Summer 1980. "The Influence of Student, Course, and Instructor Characteristics in Evaluations of University Teaching." *American Educational Research Journal* 17: 219–237.

Marsh, Herbert W. and Cooper, T.L. January 1981. "Prior Subject Interest, Students' Evaluations, and Instructional Effectiveness." *Multivariate Behavioral Research* 16: 82–104.

Marsh, Herbert W. and Overall, Jesse U. Spring 1981. "The Relative Influence of Course Level, Course Type, and Instructor on Students' Evaluations of College Teaching." *American Educational Research Journal* 18: 103–112.

Marsh, Herbert W. and Overall, Jesse U. August 1980. "Validity of Students' Evaluations of Teaching Effectiveness: Cognitive and Affective Criteria." *Journal of Educational Psychology* 72: 468–475.

Marsh, Herbert W. and Ware, John E. February 1982. "Effects of Expressiveness, Content Coverage and Incentive on Multidimensional Student Rating Scales: New Interpretations of the Dr. Fox Effect." *Journal of Educational Psychology* 74: 126–134.

Marsh, Herbert W.; Overall, Jesse U.; and Kesler, Steven P. April 1979. "Validity of Student Evaluations of Instructional Effectiveness: A Comparison of Faculty Self-Evaluations and Evaluations by Their Students." *Journal of Educational Psychology* 71: 149–160.

McKeachie, Wilbert J. October 1979. "Student Ratings of Faculty: A Reprise." *Academe* 65: 384–397.

McKeachie, Wilbert J. and Lin, Y-G. 1975. "Study I: Using Student Ratings to Improve Teaching." Ann Arbor, Michigan: University of Michigan, 1975. ED 104 284. MF-$1.17; PC-$3.65.

McKeachie, Wilbert J.; Lin, Y-G; Daugherty, Monica; Moffett, Mary M.; Neigler, Cynthia; Nork, J.; Walz, Martina; and Baldwin, T. June 1980. "Using Student Ratings and Consultation to Improve Instruction." *British Journal of Educational Psychology* 50: 168–174.

Millman, Jason, ed. 1981. *Handbook of Teacher Evaluation*. Beverly Hills, California: Sage Publications.

Murray, Deborah and Abrami, Philip C. April 1981. "The Relationship between Student/Teacher Attitude Similarity and Ratings of Instructional Effectiveness." Los Angeles: Paper presented at the annual meeting of the American Educational Research Association. ED 201 274. MF-$1.17; PC-$5.39.

Naftulin, Donald H.; Ware, John E.; and Donnelly, Frank A. July 1973. "The Doctor Fox Lecture: A Paradigm of Educational Seduction." *Journal of Medical Education* 48: 630–635.

Overall, Jesse U. and Marsh, Herbert W. June 1980. "Students' Evaluations of Instruction: A Longitudinal Study of Their Stability." *Journal of Educational Psychology* 72: 321–325.

Overall, Jesse U. and Marsh, Herbert W. December 1979. "Midterm Feedback from Students: Its Relationship to Instructional Improvement and Students' Cognitive and Affective Outcomes." *Journal of Educational Psychology* 71: 856–865.

Perry, Raymond P.; Abrami, Philip C.; Leventhal, Les; and Check, James. December 1979. "Instructor Reputation: An Expectancy Relationship Involving Student Ratings and Achievement." *Journal of Educational Psychology* 71: 776–787.

Rotem, Arie and Glasman, Naftaly S. Summer 1979. "On the Effectiveness of Students' Evaluative Feedback to University Instructors." *Review of Educational Research* 49: 497–511.

Stumpf, Stephen A.; Freedman, Richard D.; and Aguanno, Joseph C. 1979. "A Path Analysis of Factors Often Found to be Related in Student Ratings of Teaching Effectiveness." *Research in Higher Education* 11: 111–123.

Ware, John E. and Williams, Reed G. 1980. "A Reanalysis of the Doctor Fox Experiments." *Instructional Evaluation* 4: 15–18.

Ware, John E. and Williams Reed G. 1979. "Seeing Through the Dr. Fox Effect: A Response to Frey." *Instructional Evaluation* 3: 6–10.